



Tecnológico de Monterrey

Analítica de datos y herramientas de inteligencia artificial

Actividad 5 (Evaluación)

José Ignacio Hernández Rodríguez A01703130

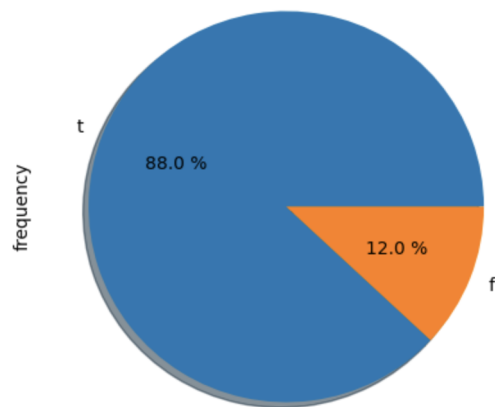
Tecnológico de Monterrey, Campus Querétaro

28/04/2023

El presente reporte aterriza en las conclusiones obtenidas de la comparación de 5 variables que se ubican en los datasets de California, Girona y DF. Las variables seleccionadas para la comparación de información entre las bases de datos son las siguientes: "host_identity_verified", "host_acceptance_rate", "host_response_time", "property_type", "number_of_reviews". Es importante mencionar que para realizar el análisis univariado de las variables es necesario que los datos de cada variable sean categóricos. De no ser así se requiere realizar una transformación de los mismos para que pasen de numéricos a categóricos, esto mediante una transformación de int a str. Este caso se presentó en la última variable.

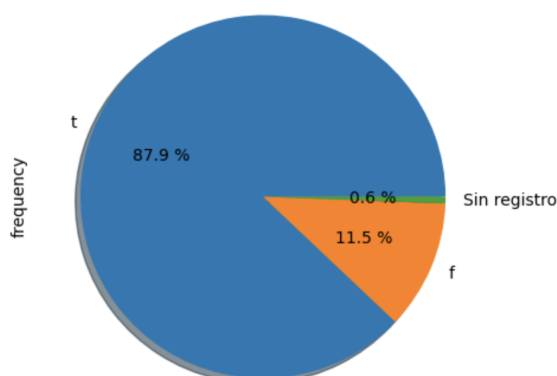
host_identity_verified

California



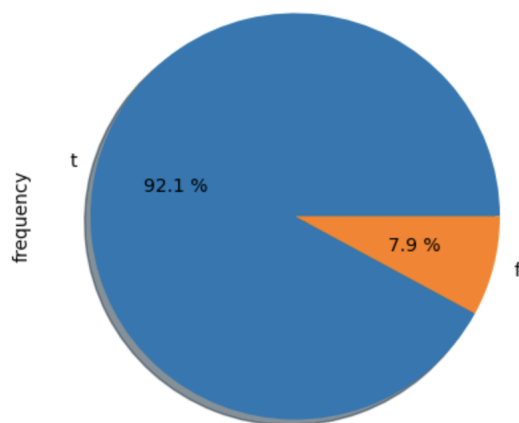
Como podemos observar, en la gráfica del dataset de California, las variables se clasifican en t (true) f (false). Claramente se puede interpretar que el 88% de los host de airbnb cuenta con su identidad verificada. Este dato representa una gran diferencia a comparación de los host que no cuentan con la verificación de su identidad.

Girona



Como podemos observar, en la gráfica del dataset de Girona, las variables se clasifican en t (true) f (false). Claramente se puede interpretar que el 87.9% de los host de airbnb cuenta con su identidad verificada. Este dato representa una gran diferencia a comparación de los host que no cuentan con la verificación de su identidad. Una peculiaridad que se presenta en esta gráfica es que se identifica un 0.6% de host que no cuentan con registro alguno en esta variable

DF

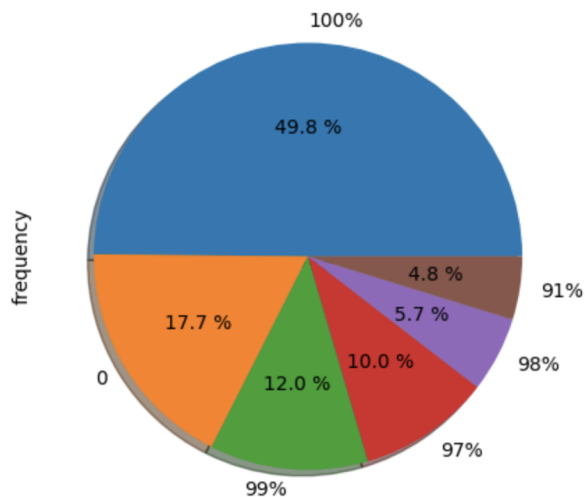


Como podemos observar, en la gráfica del dataset de DF, las variables se clasifican en t (true) f (false). Claramente se puede interpretar que el 92.1% de los host de airbnb cuenta con su identidad verificada. Este dato representa una gran diferencia a comparación de los host que no cuentan con la verificación de su identidad, cuyo porcentaje es apenas de 7.9%.

En conclusión, se puede decir que el país con mayor cantidad de hosts con la identidad verificada es México DF. A pesar de ser la ciudad con mejores resultados, el par de ciudades restantes, se ubican ligeramente por debajo. La ciudad con datos con menor veracidad es Girona ya que cuenta con un pequeño porcentaje de usuarios sin registros.

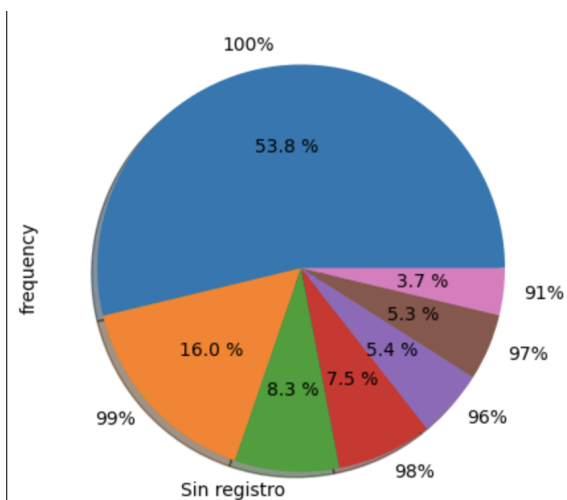
host_acceptance_rate

California



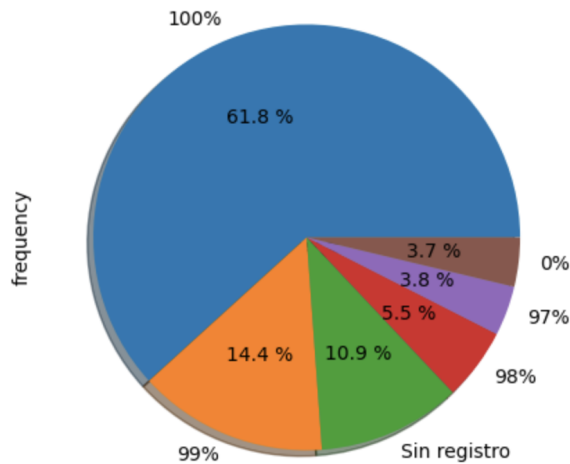
Como podemos observar, en la gráfica se clasifican los porcentajes de aceptación de cada host. En esta gráfica se presentan un total de 6 porcentajes, mismos que presentaron una mayor frecuencia a comparación de los demás porcentajes. Con el 49.8% se encuentran los host que presentan una aceptación del 100%. Esto indica que en este país es muy probable que el host acepte una solicitud.

Girona



Como podemos observar, en la gráfica se clasifican los porcentajes de aceptación de cada host. En esta gráfica se presentan un total de 7 porcentajes, mismos que presentaron una mayor frecuencia a comparación de los demás porcentajes. Con el 53.8% se encuentran los host que presentan una aceptación del 100%. Esto indica que en este país es muy probable que el host acepte una solicitud. Un aspecto importante a considerar es que en este país el número de hosts sin registro en esta categoría ya se torna significativo, pues el 8.3% no cuenta con registro alguno. Por lo que este porcentaje puede alternar las medidas de dispersión de la variable.

DF



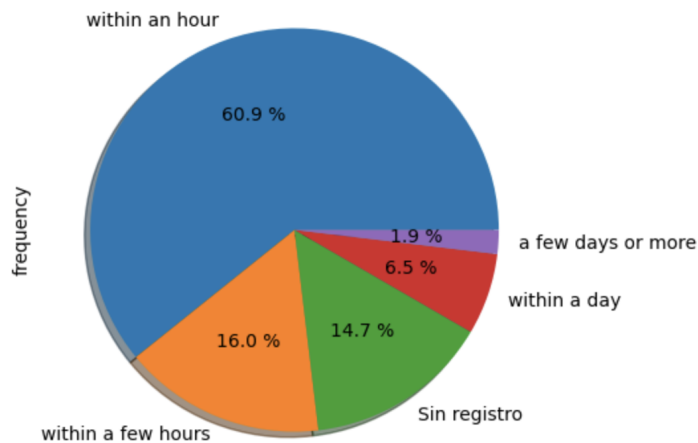
Como podemos observar, en la gráfica se clasifican los porcentajes de aceptación de cada host. En esta gráfica se presentan un total de 6 porcentajes, mismos que presentaron una mayor frecuencia a comparación de los demás porcentajes. Con el 61.8% se encuentran los host que presentan una aceptación del 100%. Esto indica que en este país es muy probable que el host acepte una solicitud. Un aspecto importante a considerar es que en este país el número de hosts sin registro en esta categoría ya se torna significativo, pues el 10.9% no cuenta con registro alguno. Por lo que este porcentaje puede alternar las medidas de dispersión de la variable. A diferencia de las demás ciudades, esta ciudad presenta una aceptación del 0% muy baja, pues solamente el 3.7% de los hosts cuentan con este registro.

Conclusión

En conclusión, se puede decir que la ciudad con una mayor tasa de aceptación por parte de los hosts es DF. Esto está sustentado con el 61.85% de hosts que cuentan con el 100% de aceptación. Mientras que la ciudad con una menor tasa de aceptación es California, ya que el 0% representa el 17.7% del total de dicha ciudad.

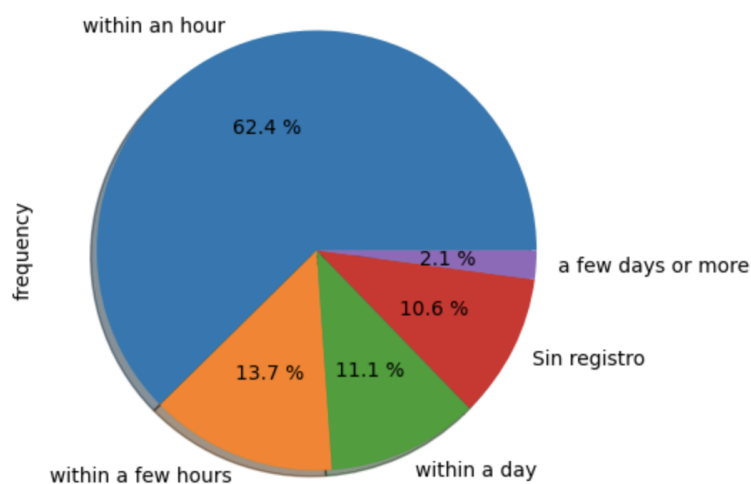
host_response_time

California



En la gráfica se presenta el tiempo de respuesta de los hosts, en esta gráfica se presentan 5 categorías. Como podemos observar, en esta gráfica los hosts presentan una gran capacidad de respuesta en cuanto al tiempo, pues solamente tardan en su mayoría aproximadamente 1 hora, esto se representa con el 60.9%.

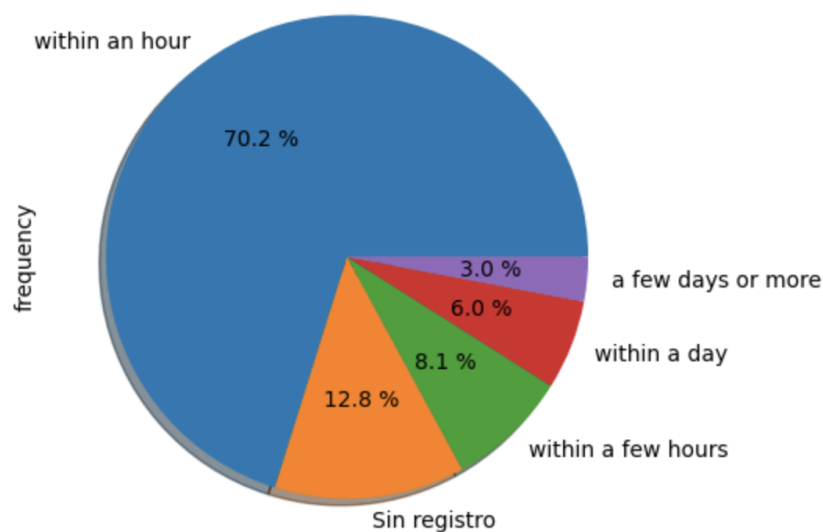
Girona



En la gráfica se presenta el tiempo de respuesta de los hosts, en esta gráfica se presentan 5 categorías. Como podemos observar, en esta gráfica los hosts presentan una gran capacidad de respuesta en cuanto al tiempo, pues solamente tardan en su mayoría aproximadamente 1 hora, esto se representa con el 62.4%.

Al igual que la ciudad anterior, se presentan hosts sin registro, sin embargo en Girona, la cantidad de hosts sin registro en esta variable va en aumento. Llegando así a 10.6%.

DF



En la gráfica se presenta el tiempo de respuesta de los hosts, en esta gráfica se presentan 5 categorías. Como podemos observar, en esta gráfica los hosts presentan una gran capacidad de respuesta en cuanto al tiempo, pues solamente tardan en su mayoría aproximadamente 1 hora, esto se representa con el 70.2%.

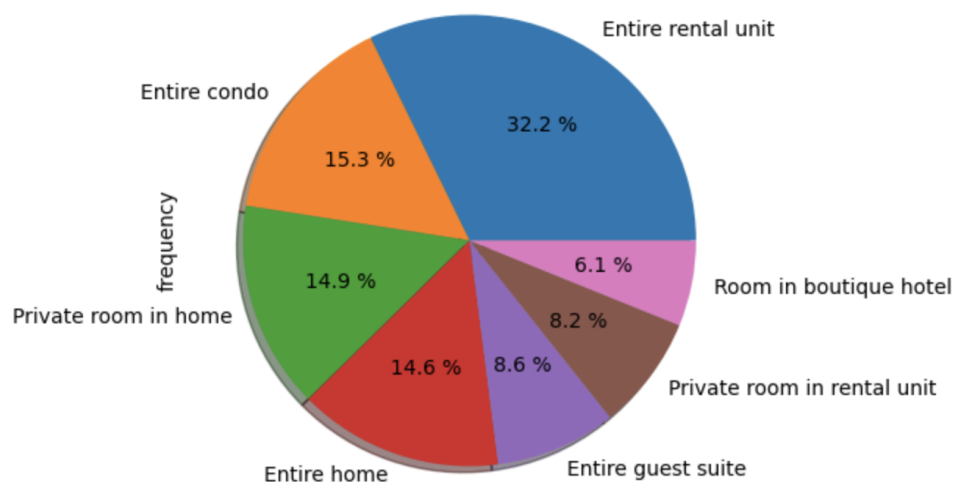
Al igual que la ciudad anterior, se presentan hosts sin registro, sin embargo en DF, la cantidad de hosts sin registro en esta variable va en aumento llegando a 12.8%.

Conclusión

En conclusión, se puede decir que la ciudad con un mejor tiempo de respuesta es DF. El par de ciudades restantes, se acercan a DF, sin embargo se encuentran ligeramente por debajo. A pesar de que DF es la ciudad con un mejor tiempo de respuesta, también es la ciudad con mayores datos de registro. Lo que podría llegar a modificar el sentido de la información final en caso de que estos datos se completaran.

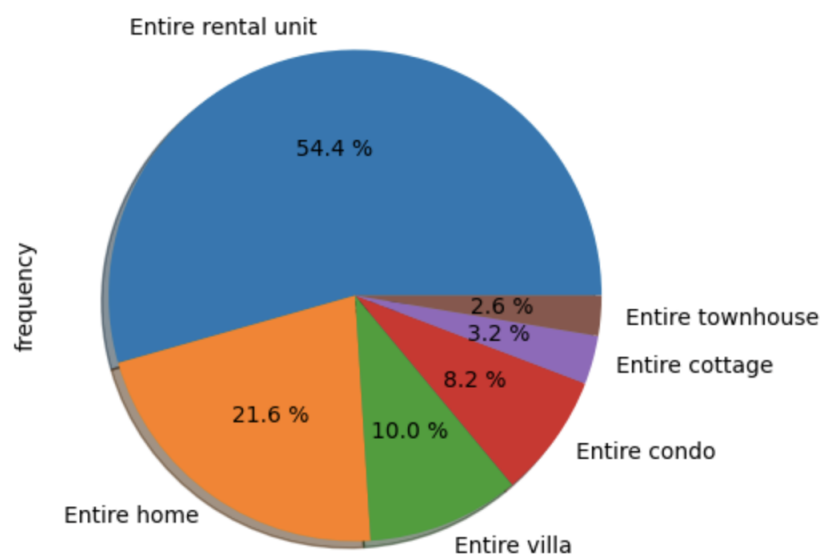
property_type

California



Como podemos observar, en la gráfica se clasifican los porcentajes del tipo de propiedad en la plataforma. En esta gráfica se presentan un total de 7 porcentajes, mismos que presentaron una mayor frecuencia a comparación de los demás porcentajes. Con el 32.2% se encuentra el tipo de entire rental unit, esto quiere decir que se renta la unidad en su totalidad. Seguido de esta categoría, se ubica entire condo con 15.3%. Es importante mencionar que el resto de tipo de propiedades presentan porcentajes menores, sin embargo no existe una gran diferencia entre los mismos.

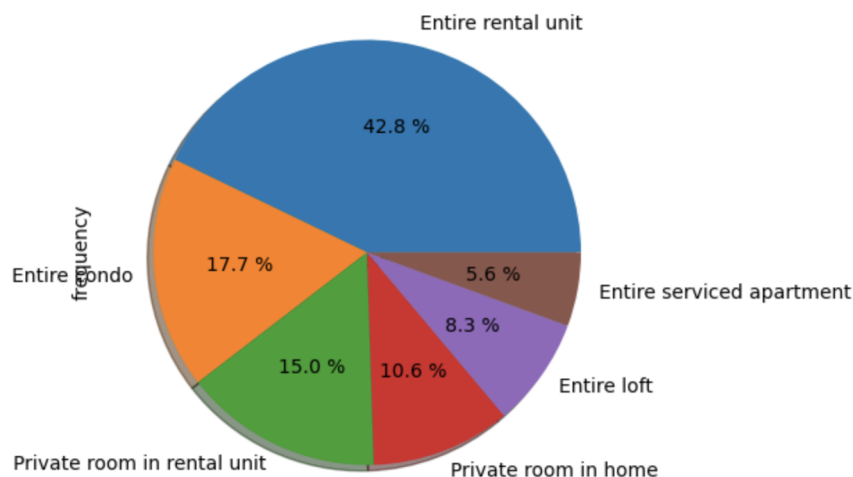
Girona



Como podemos observar, en la gráfica se clasifican los porcentajes del tipo de propiedad en la plataforma para la ciudad de Girona. En esta gráfica se presentan

un total de 6 porcentajes, mismos que presentaron una mayor frecuencia a comparación de los demás porcentajes. Con el 54.4% se encuentra el tipo de entire rental unit, esto quiere decir que se renta la unidad en su totalidad. Seguido de esta categoría, se ubica entire home con 15.3%. Es importante mencionar que el resto de tipo de propiedades presentan porcentajes menores, sin embargo no existe una gran diferencia entre los mismos.

DF



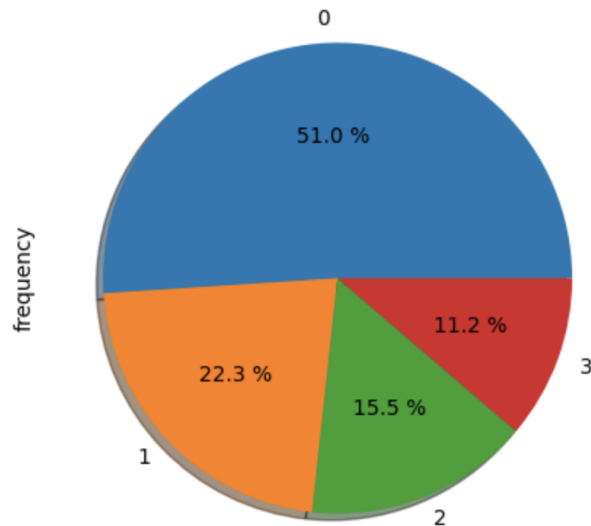
Como podemos observar, en la gráfica se clasifican los porcentajes del tipo de propiedad en la plataforma. En esta gráfica se presentan un total de 6 porcentajes, mismos que presentaron una mayor frecuencia a comparación de los demás porcentajes. Con el 42.8% se encuentra el tipo de entire rental unit, esto quiere decir que se renta la unidad en su totalidad. Seguido de esta categoría, se ubica entire condo con 17.7%. Seguido de private room in rental unit con 15.0%.

Conclusión

A diferencia de las gráficas de las variables anteriores. En las gráficas de la variable property type no se presentan valores sin registro. Esto habla de la veracidad de la información y que no se presentarán variaciones en los datos. En conclusión, se puede decir que el tipo de propiedad que más predomina en las tres ciudades es el entire rental unit.

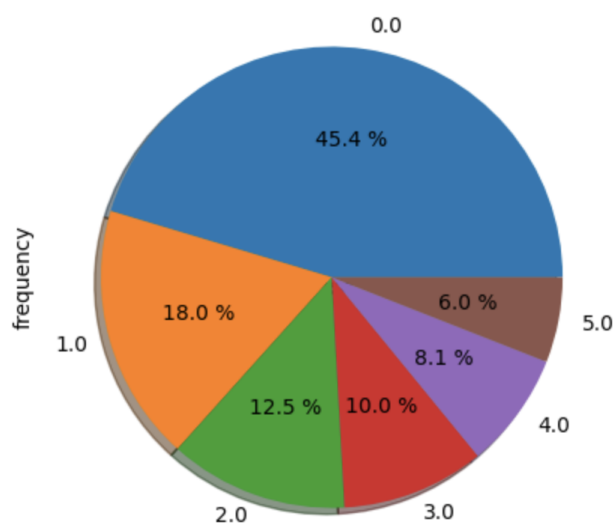
number_of_reviews

California



Como podemos ver, en la gráfica se presentan las frecuencias de las reseñas que reciben los hosts sobre sus servicios. En esta gráfica se presentan 4 categorías numéricas sobre las reseñas. La categoría que predomina es aquella con 0 reseñas. Esto indica que los huéspedes no tienen mucho interés en dejar su reseña. Seguido de este porcentaje se ubica la categoría de 1 reseña la cuál cuenta con 22.3%. Esta categoría representa aproximadamente la suma de las siguientes dos categorías (2 y 3 reseñas).

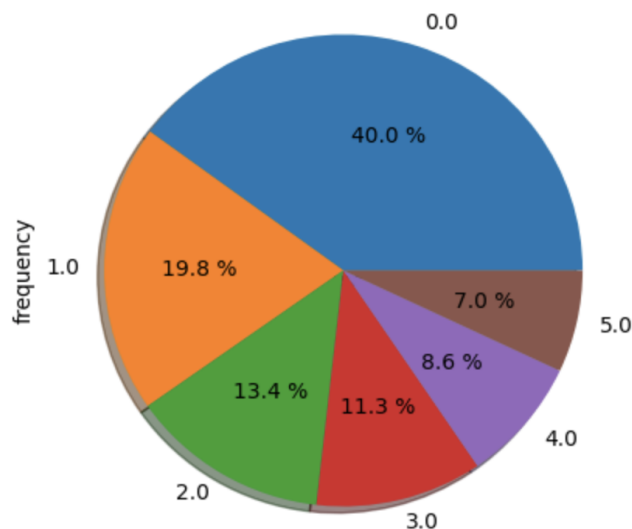
Girona



Como podemos ver, en la gráfica se presentan las frecuencias de las reseñas que reciben los hosts sobre sus servicios. En esta gráfica se presentan 6 categorías numéricas sobre las reseñas. La categoría que predomina es aquella con 0 reseñas con 45.4%. Esto indica que los huéspedes no tienen mucho interés en dejar su

reseña. Seguido de este porcentaje se ubica la categoría de 1 reseña la cuál cuenta con 18.0%. En esta base de datos de Girona, podemos observar que existen más categorías de reseñas.

DF



Como podemos ver, en la gráfica se presentan las frecuencias de las reseñas que reciben los hosts sobre sus servicios. En esta gráfica se presentan 6 categorías numéricas sobre las reseñas. La categoría que predomina es aquella con 0 reseñas con 40.0%. Esto indica que los huéspedes no tienen mucho interés en dejar su reseña. Seguido de este porcentaje se ubica la categoría de 1 reseña la cuál cuenta con 19.8%. En esta base de datos de DF, podemos observar que existen más categorías de reseñas. Con mayores porcentajes en comparación con las otras ciudades.

Conclusión

Como conclusión se puede decir que comparando las gráficas de las tres ciudades, la ciudad con un mayor número de reviews es Girona, con un 7% que hace referencia a 5 reseñas. Mientras que la ciudad con menos reseñas es California con un 50% que hace referencia a 0 reseñas. Tomando en cuenta los resultados, se puede interpretar que los huéspedes no dan una gran importancia a las reseñas, por lo que las omiten en muchas ocasiones.